

## SEQUENCE LISTING

<110> Sticklen, Masomeh B  
 Maqbool, Shahina B  
 Dale, Bruce E

<120> TRANSGENIC PLANTS CONTAINING LIGNINASE AND CELLULASE WHICH  
 DEGRADE  
 LIGNIN AND CELLULOSE TO FERMENTABLE SUGARS

<130> MSU 4.1-539

<150> 60/242,408

<151> 2000-10-20

<160> 19

<170> PatentIn version 3.1

<210> 1

<211> 1110

<212> DNA

<213> Oryza sativa

<400> 1

```

gggtcggaga tgccaccacg gccacaaccc acgagcccgg cgcgacacca ccgcgcgcgt      60
tgagccagcc acaaacgccc gcggataggc gcgccgcacg cggccaatcc taccacatcc      120
ccggcctcgg cggtcgagc gccgtgccat ccgatccgct gagttttggc tattttatacg      180
taccgcggga gcctgtgtgc agagagtgc tctcaagaag tactcgagca aagaaggaga      240
gagcttggtg agctgcagag atggccccct ccgtgatggc gtcgtcggcc accaccgtcg      300
ctcccttcca gggctcaagt ccaccgccgg catgccgtcg ccgcgcgtcc gaactccagc      360
ttcggcaacg tcagcatggc ggcaggatca ggtgcatgca ggtaattacc tactgatcca      420
acacacattc ttcttcttct tcttcttctt aaccaacatt aaccaacaac tcaattatcg      480
tttattcatt gaggtgtggc cgattgaggg catcaagaag ttcgagacct tctctacct      540
gccaccgctc accgtggagg acctcctgaa gcagatcgag tacctagctc cgttccaagt      600
ggtgccctgc ctcgagttca gcaaggctcg atttgcttac cgtgagaacc acaagtcccc      660
tggtacttac gacggcaggt actggaccat gtggaagctg cccatgttcg ggtgcaccga      720
cgccaccag gtcgtcaagg agctcgagga ggccaagaag gcgtaccctg atgcattcgt      780
ccgtatcatc ggcttcgaca acgttaggca ggtgcagctc atcagcttca tcgcctacaa      840
cccgggctgc gaggagtctg gtggcaacta agccgtcatc gtcatatata gcctcgttta      900
attgttcac tctgattcga tgatgtctcc caccttgttt cgtgtgttcc cagtttgttt      960
catcgtcttt tgattttacc ggccgtgctc tgcttttggt ttttcttttc acctgattct     1020
ctctctgact tgatgtaaga gtggtatctg ctacgactat atgttggttg ggtgaggcat     1080

```

atgtgaatga aatctatgaa agctccggct

1110

<210> 2  
<211> 38  
<212> PRT  
<213> Oryza sativa

<400> 2

Met Ala Pro Ser Val Met Ala Ser Ser Ala Thr Thr Val Ala Pro Phe  
1 5 10 15

Gln Gly Ser Ser Pro Pro Pro Ala Cys Arg Arg Pro Pro Ser Glu Leu  
20 25 30

Gln Leu Arg Gln Arg Gln  
35

<210> 3  
<211> 6  
<212> PRT  
<213> synthetic peptide

<220>  
<221> SIGNAL  
<222> (1)..(6)  
<223> targets the peroxisomes of plants

<400> 3

Arg Ala Val Ala Arg Leu  
1 5

<210> 4  
<211> 3004  
<212> DNA  
<213> Acidothermus cellulolyticus

<220>  
<221> CDS  
<222> (824)..(2512)  
<223> E I beta-1,4-endoglucanase precursor

<400> 4

ggatccacgt tgtacaaggt cacctgtccg tcgttctggt agagcggcgg gatggtcacc 60  
cgcacgatct ctcttttggt gatgtcgacg gtcacgtggt tacggtttgc ctcggccgcg 120  
attttcgcgc tcgggcttgc tccggctgtc gggttcgggt tggcgtggtg tgcggagcac 180  
gccgaggcga tcccaatgag ggcaaggga agagcggagc cgatggcacg tcgggtggcc 240  
gatggggtac gccgatgggg cgtggcgctcc ccgccgcgga cagaaccgga tgcggaatag 300  
gtcacggtgc gacatggtgc cgtaccgcgg acccgatga caagggtggg tgcgcgggtc 360  
gcctgtgagc tgccggctgg cgtctggatc atgggaacga tcccaccatt ccccgcaatc 420

gacgcgagtcg	ggagcagggc	ggcgcgagcc	ggaccgtgtg	gtcgcagccgg	acgattcgcc		480										
catacggtgc	tgcaatgcc	agcgccatgt	tgtcaatccg	ccaaatgcag	caatgcacac		540										
atggacaggg	attgtgactc	tgagtaatga	ttggattgcc	ttcttgccgc	ctacgcgtta		600										
cgcagagtag	gcgactgtat	gcggtagggt	ggcgctccag	ccgtgggctg	gacatgcctg		660										
ctgcgaactc	ttgacacgtc	tggttgaacg	cgcaatactc	ccaacaccga	tgggatcggt		720										
cccataagtt	tccgtctcac	aacagaatcg	gtgcgcctc	atgatcaacg	tgaaaggagt		780										
acggggggaga	acagacgggg	gagaaaccaa	cggggggattg	gcg	gtg	ccg	cg	gca									835
					Val	Pro	Arg	Ala									
ttg	cg	cga	gtg	cct	ggc	tcg	cg	gtg	atg	ctg	cg	gtc	ggc	gtc	gtc		883
Leu	Arg	Arg	Val	Pro	Gly	Ser	Arg	Val	Met	Leu	Arg	Val	Gly	Val	Val		
5					10					15					20		
gtc	gc	gtg	ctg	gca	ttg	gtt	gcc	gca	ctc	gcc	aac	cta	gcc	gtg	ccg		931
Val	Ala	Val	Leu	Ala	Leu	Val	Ala	Ala	Leu	Ala	Asn	Leu	Ala	Val	Pro		
				25					30					35			
cg	cc	gct	cg	gcc	gc	ggc	ggc	ggc	tat	tgg	cac	acg	agc	ggc	cg		979
Arg	Pro	Ala	Arg	Ala	Ala	Gly	Gly	Gly	Tyr	Trp	His	Thr	Ser	Gly	Arg		
			40					45					50				
gag	atc	ctg	gac	gc	aac	aac	gtg	ccg	gta	cg	atc	gcc	ggc	atc	aac		1027
Glu	Ile	Leu	Asp	Ala	Asn	Asn	Val	Pro	Val	Arg	Ile	Ala	Gly	Ile	Asn		
		55					60					65					
tgg	ttt	ggg	ttc	gaa	acc	tgc	aat	tac	gtc	gtg	cac	ggt	ctc	tgg	tca		1075
Trp	Phe	Gly	Phe	Glu	Thr	Cys	Asn	Tyr	Val	Val	His	Gly	Leu	Trp	Ser		
	70					75					80						
cg	gac	tac	cg	agc	atg	ctc	gac	cag	ata	aag	tcg	ctc	ggc	tac	aac		1123
Arg	Asp	Tyr	Arg	Ser	Met	Leu	Asp	Gln	Ile	Lys	Ser	Leu	Gly	Tyr	Asn		
85				90						95					100		
aca	atc	cg	ctg	ccg	tac	tct	gac	gac	att	ctc	aag	ccg	ggc	acc	atg		1171
Thr	Ile	Arg	Leu	Pro	Tyr	Ser	Asp	Asp	Ile	Leu	Lys	Pro	Gly	Thr	Met		
				105					110					115			
ccg	aac	agc	atc	aat	ttt	tac	cag	atg	aat	cag	gac	ctg	cag	ggt	ctg		1219
Pro	Asn	Ser	Ile	Asn	Phe	Tyr	Gln	Met	Asn	Gln	Asp	Leu	Gln	Gly	Leu		
			120					125					130				
acg	tcc	ttg	cag	gtc	atg	gac	aaa	atc	gtc	gc	tac	gcc	ggt	cag	atc		1267
Thr	Ser	Leu	Gln	Val	Met	Asp	Lys	Ile	Val	Ala	Tyr	Ala	Gly	Gln	Ile		
		135					140					145					
ggc	ctg	cg	atc	att	ctt	gac	cg	cac	cga	ccg	gat	tgc	agc	ggg	cag		1315
Gly	Leu	Arg	Ile	Ile	Leu	Asp	Arg	His	Arg	Pro	Asp	Cys	Ser	Gly	Gln		
	150					155					160						
tcg	gc	ctg	tgg	tac	acg	agc	agc	gtc	tcg	gag	gct	acg	tgg	att	tcc		1363
Ser	Ala	Leu															

ggc ttt gac ttg cac aac gag ccg cat gac ccg gcc tgc tgg ggc tgc	1459
Gly Phe Asp Leu His Asn Glu Pro His Asp Pro Ala Cys Trp Gly Cys	
200 205 210	
ggc gat ccg agc atc gac tgg cga ttg gcc gcc gag cgg gcc gga aac	1507
Gly Asp Pro Ser Ile Asp Trp Arg Leu Ala Ala Glu Arg Ala Gly Asn	
215 220 225	
gcc gtg ctc tcg gtg aat ccg aac ctg ctc att ttc gtc gaa ggt gtg	1555
Ala Val Leu Ser Val Asn Pro Asn Leu Leu Ile Phe Val Glu Gly Val	
230 235 240	
cag agc tac aac gga gac tcc tac tgg tgg ggc gcc aac ctg caa gga	1603
Gln Ser Tyr Asn Gly Asp Ser Tyr Trp Trp Gly Gly Asn Leu Gln Gly	
245 250 255 260	
gcc ggc cag tac ccg gtc gtg ctg aac gtg ccg aac cgc ctg gtg tac	1651
Ala Gly Gln Tyr Val Val Leu Asn Val Pro Asn Arg Leu Val Tyr	
265 270 275	
tcg gcg cac gac tac gcg acg agc gtc tac ccg cag acg tgg ttc agc	1699
Ser Ala His Asp Tyr Ala Thr Ser Val Tyr Pro Gln Thr Trp Phe Ser	
280 285 290	
gat ccg acc ttc ccc aac aac atg ccc ggc atc tgg aac aag aac tgg	1747
Asp Pro Thr Phe Pro Asn Asn Met Pro Gly Ile Trp Asn Lys Asn Trp	
295 300 305	
gga tac ctc ttc aat cag aac att gca ccg gta tgg ctg gcc gaa ttc	1795
Gly Tyr Leu Phe Asn Gln Asn Ile Ala Pro Val Trp Leu Gly Glu Phe	
310 315 320	
ggt acg aca ctg caa tcc acg acc gac cag acg tgg ctg aag acg ctc	1843
Gly Thr Thr Leu Gln Ser Thr Thr Asp Gln Thr Trp Leu Lys Thr Leu	
325 330 335 340	
gtc cag tac cta cgg ccg acc gcg caa tac ggt gcg gac agc ttc cag	1891
Val Gln Tyr Leu Arg Pro Thr Ala Gln Tyr Gly Ala Asp Ser Phe Gln	
345 350 355	
tgg acc ttc tgg tcc tgg aac ccc gat tcc ggc gac aca gga gga att	1939
Trp Thr Phe Trp Ser Trp Asn Pro Asp Ser Gly Asp Thr Gly Gly Ile	
360 365 370	
ctc aag gat gac tgg cag acg gtc gac aca gta aaa gac ggc tat ctc	1987
Leu Lys Asp Asp Trp Gln Thr Val Asp Thr Val Lys Asp Gly Tyr Leu	
375 380 385	
gcg ccg atc aag tcg tcg att ttc gat cct gtc ggc gcg tct gca tcg	2035
Ala Pro Ile Lys Ser Ser Ile Phe Asp Pro Val Gly Ala Ser Ala Ser	
390 395 400	
cct agc agt caa ccg tcc ccg tcg gtg tcg ccg tct ccg tcg ccg agc	2083
Pro Ser Ser Gln Pro Ser Pro Ser Val Ser Pro Ser Pro Ser Pro Ser	
405 410 415 420	
ccg tcg gcg agt cgg acg ccg acg cct act ccg acg ccg aca gcc agc	2131
Pro Ser Ala Ser Arg Thr Pro Thr Pro Thr Pro Thr Pro Thr Ala Ser	
425 430 435	
ccg acg cca acg ctg acc cct act gct acg ccc acg ccc acg gca agc	2179
Pro Thr Pro Thr Leu Thr Pro Thr Ala Thr Pro Thr Pro Thr Ala Ser	
440 445 450	

ccg acg ccg tca ccg acg gca gcc tcc gga gcc cgc tgc acc gcg agt	2227
Pro Thr Pro Ser Pro Thr Ala Ala Ser Gly Ala Arg Cys Thr Ala Ser	
455 460 465	
tac cag gtc aac agc gat tgg ggc aat ggc ttc acg gta acg gtg gcc	2275
Tyr Gln Val Asn Ser Asp Trp Gly Asn Gly Phe Thr Val Thr Val Ala	
470 475 480	
gtg aca aat tcc gga tcc gtc gcg acc aag aca tgg acg gtc agt tgg	2323
Val Thr Asn Ser Gly Ser Val Ala Thr Lys Thr Trp Thr Val Ser Trp	
485 490 495 500	
aca ttc ggc gga aat cag acg att acc aat tcg tgg aat gca gcg gtc	2371
Thr Phe Gly Gly Asn Gln Thr Ile Thr Asn Ser Trp Asn Ala Ala Val	
505 510 515	
acg cag aac ggt cag tcg gta acg gct cgg aat atg agt tat aac aac	2419
Thr Gln Asn Gly Gln Ser Val Thr Ala Arg Asn Met Ser Tyr Asn Asn	
520 525 530	
gtg att cag cct ggt cag aac acc acg ttc gga ttc cag gcg agc tat	2467
Val Ile Gln Pro Gly Gln Asn Thr Thr Phe Gly Phe Gln Ala Ser Tyr	
535 540 545	
acc gga agc aac gcg gca ccg aca gtc gcc tgc gca gca agt taa	2512
Thr Gly Ser Asn Ala Ala Pro Thr Val Ala Cys Ala Ala Ser	
550 555 560	
tacgtcgggg agccgacggg aggggtccgga ccgtcgggttc cccggcttcc acctatggag	2572
cgaacccaac aatccggacg gaactgcagg taccagagag gaacgacacg aatgcccgcc	2632
atctcaaaac ggctgcgagc cggcgctctc gccggggcg tgcagcatcg agcctccatc	2692
gtgccgctgg cgatgcagca tcttgccatc gccgcgacgc acgtcgacaa tccctatgcg	2752
ggagcgacct tcttcgtcaa cccgtactgg gcgcaagaag tacagagcga acggcgaacc	2812
agaccaatgc cactctcgca gcgaaaatgc gcgtcgtttc cacatattcg acggccgtct	2872
ggatggaccg catcgctgcg atcaacggcg tcaacggcg acccggttg acgacatc	2932
tggacgccgc cctctccag cagcagggaa ccaccctga agtcattgag attgtcatct	2992
acgatctgcc gg	3004

<210> 5  
 <211> 562  
 <212> PRT  
 <213> Acidothermus cellulolyticus

<400> 5

Val Pro Arg Ala Leu Arg Arg Val Pro Gly Ser Arg Val Met Leu Arg
1 5 10 15

Val Gly Val Val Val Ala Val Leu Ala Leu Val Ala Ala Leu Ala Asn
20 25 30

Leu Ala Val Pro Arg Pro Ala Arg Ala Ala Gly Gly Gly Tyr Trp His
35 40 45

Thr Ser Gly Arg Glu Ile Leu Asp Ala Asn Asn Val Pro Val Arg Ile  
 50 55 60

Ala Gly Ile Asn Trp Phe Gly Phe Glu Thr Cys Asn Tyr Val Val His  
 65 70 75 80

Gly Leu Trp Ser Arg Asp Tyr Arg Ser Met Leu Asp Gln Ile Lys Ser  
 85 90 95

Leu Gly Tyr Asn Thr Ile Arg Leu Pro Tyr Ser Asp Asp Ile Leu Lys  
 100 105 110

Pro Gly Thr Met Pro Asn Ser Ile Asn Phe Tyr Gln Met Asn Gln Asp  
 115 120 125

Leu Gln Gly Leu Thr Ser Leu Gln Val Met Asp Lys Ile Val Ala Tyr  
 130 135 140

Ala Gly Gln Ile Gly Leu Arg Ile Ile Leu Asp Arg His Arg Pro Asp  
 145 150 155 160

Cys Ser Gly Gln Ser Ala Leu Trp Tyr Thr Ser Ser Val Ser Glu Ala  
 165 170 175

Thr Trp Ile Ser Asp Leu Gln Ala Leu Ala Gln Arg Tyr Lys Gly Asn  
 180 185 190

Pro Thr Val Val Gly Phe Asp Leu His Asn Glu Pro His Asp Pro Ala  
 195 200 205

Cys Trp Gly Cys Gly Asp Pro Ser Ile Asp Trp Arg Leu Ala Ala Glu  
 210 215 220

Arg Ala Gly Asn Ala Val Leu Ser Val Asn Pro Asn Leu Leu Ile Phe  
 225 230 235 240

Val Glu Gly Val Gln Ser Tyr Asn Gly Asp Ser Tyr Trp Trp Gly Gly  
 245 250 255

Asn Leu Gln Gly Ala Gly Gln Tyr Pro Val Val Leu Asn Val Pro Asn  
 260 265 270

Arg Leu Val Tyr Ser Ala His Asp Tyr Ala Thr Ser Val Tyr Pro Gln  
 275 280 285

Thr Trp Phe Ser Asp Pro Thr Phe Pro Asn Asn Met Pro Gly Ile Trp  
 290 295 300

Asn Lys Asn Trp Gly Tyr Leu Phe Asn Gln Asn Ile Ala Pro Val Trp  
305 310 315 320

Leu Gly Glu Phe Gly Thr Thr Leu Gln Ser Thr Thr Asp Gln Thr Trp  
325 330 335

Leu Lys Thr Leu Val Gln Tyr Leu Arg Pro Thr Ala Gln Tyr Gly Ala  
340 345 350

Asp Ser Phe Gln Trp Thr Phe Trp Ser Trp Asn Pro Asp Ser Gly Asp  
355 360 365

Thr Gly Gly Ile Leu Lys Asp Asp Trp Gln Thr Val Asp Thr Val Lys  
370 375 380

Asp Gly Tyr Leu Ala Pro Ile Lys Ser Ser Ile Phe Asp Pro Val Gly  
385 390 395 400

Ala Ser Ala Ser Pro Ser Ser Gln Pro Ser Pro Ser Val Ser Pro Ser  
405 410 415

Pro Ser Pro Ser Pro Ser Ala Ser Arg Thr Pro Thr Pro Thr Pro Thr  
420 425 430

Pro Thr Ala Ser Pro Thr Pro Thr Leu Thr Pro Thr Ala Thr Pro Thr  
435 440 445

Pro Thr Ala Ser Pro Thr Pro Ser Pro Thr Ala Ala Ser Gly Ala Arg  
450 455 460

Cys Thr Ala Ser Tyr Gln Val Asn Ser Asp Trp Gly Asn Gly Phe Thr  
465 470 475 480

Val Thr Val Ala Val Thr Asn Ser Gly Ser Val Ala Thr Lys Thr Trp  
485 490 495

Thr Val Ser Trp Thr Phe Gly Gly Asn Gln Thr Ile Thr Asn Ser Trp  
500 505 510

Asn Ala Ala Val Thr Gln Asn Gly Gln Ser Val Thr Ala Arg Asn Met  
515 520 525

Ser Tyr Asn Asn Val Ile Gln Pro Gly Gln Asn Thr Thr Phe Gly Phe  
530 535 540

Gln Ala Ser Tyr Thr Gly Ser Asn Ala Ala Pro Thr Val Ala Cys Ala  
545 550 555 560

Ala Ser

<210> 6  
 <211> 1467  
 <212> DNA  
 <213> Actinomyces naeslundii

<220>  
 <221> CDS  
 <222> (1)..(1467)  
 <223> beta-glucosidase

<220>  
 <221> misc\_feature  
 <222> (339)..(339)  
 <223> nucleotide is uncertain

<220>  
 <221> misc\_feature  
 <222> (947)..(947)  
 <223> nucleotide is uncertain

<220>  
 <221> misc\_feature  
 <222> (443)..(443)  
 <223> nucleotide is uncertain

<400> 6

atg acc gcc acg tcc act act tct aag agc aat ccg aac ttc ccc gac	48
Met Thr Ala Thr Ser Thr Thr Ser Lys Ser Asn Pro Asn Phe Pro Asp	
1 5 10 15	
ggc ttc ctg tgg ggc ggg gcc acc gcc gcc aac cag atc gag ggc gct	96
Gly Phe Leu Trp Gly Gly Ala Thr Ala Ala Asn Gln Ile Glu Gly Ala	
20 25 30	
tac aac gag gac ggc aag ggc ctg tcc gtc cag gac gtc atg cct cgg	144
Tyr Asn Glu Asp Gly Lys Gly Leu Ser Val Gln Asp Val Met Pro Arg	
35 40 45	
ggc atc atg gcc cac ccc acc cag gct ccc aca ccg gat aac ctt caa	192
Gly Ile Met Ala His Pro Thr Gln Ala Pro Thr Pro Asp Asn Leu Gln	
50 55 60	
gct cga ggc gat cga cct tct acc acc gct tac gcc gag gac atc tcc	240
Ala Arg Gly Asp Arg Pro Ser Thr Thr Ala Tyr Ala Glu Asp Ile Ser	
65 70 75 80	
ctg ttc gcg gag atg ggt ttc aag gtc ttc cgc ttc tcc atc gcc tgg	288
Leu Phe Ala Glu Met Gly Phe Lys Val Phe Arg Phe Ser Ile Ala Trp	
85 90 95	
agc cgc atc ttc ccg ctc ggc gac gag acc gag ccc aat gag gaa gga	336
Ser Arg Ile Phe Pro Leu Gly Asp Glu Thr Glu Pro Asn Glu Glu Gly	
100 105 110	
ctn gcc ttc tac gac cgg gtc ctc gac gag ctc gag aag cac ggg atc	384



Xaa	Ala	Phe	Tyr	Asp	Arg	Val	Leu	Asp	Glu	Leu	Glu	Lys	His	Gly	Ile	
		115					120					125				
gag	cca	ctg	gtc	acc	atc	agc	cac	tac	gag	acc	ccg	ctg	cac	ctg	gcg	432
Glu	Pro	Leu	Val	Thr	Ile	Ser	His	Tyr	Glu	Thr	Pro	Leu	His	Leu	Ala	
	130					135					140					
cgc	acc	tac	gnc	ggc	tgg	acc	gac	cgc	cgc	ctc	atc	ggc	ttc	ttc	gag	480
Arg	Thr	Tyr	Xaa	Gly	Trp	Thr	Asp	Arg	Arg	Leu	Ile	Gly	Phe	Phe	Glu	
145					150					155					160	
cgc	tac	gcc	cgc	acc	ctg	ttc	gag	cgc	tat	ggc	aag	cgg	gtc	aag	tac	528
Arg	Tyr	Ala	Arg	Thr	Leu	Phe	Glu	Arg	Tyr	Gly	Lys	Arg	Val	Lys	Tyr	
				165					170					175		
tgg	ctc	acc	ttc	aac	gag	atc	aac	tcc	gtg	ctc	cat	gag	ccc	ttc	cta	576
Trp	Leu	Thr	Phe	Asn	Glu	Ile	Asn	Ser	Val	Leu	His	Glu	Pro	Phe	Leu	
			180					185					190			
tct	ggg	ggc	gtc	gcc	acg	ccc	aag	gac	agg	ccc	ccc	gag	cag	gac	ctc	624
Ser	Gly	Gly	Val	Ala	Thr	Pro	Lys	Asp	Arg	Pro	Pro	Glu	Gln	Asp	Leu	
	195						200					205				
tac	cag	gcc	atc	caa	aac	gag	ctc	gtc	gcc	tcc	gcg	gcc	gcg	acc	agg	672
Tyr	Gln	Ala	Ile	Gln	Asn	Glu	Leu	Val	Ala	Ser		Ala	Ala	Thr	Arg	
	210					215					220					
atc	gcc	cat	gag	acc	aac	ccc	gac	atc	cag	gtc	ggc	tgc	atg	atc	ctg	720
Ile	Ala	His	Glu	Thr	Asn	Pro	Asp	Ile	Gln	Val	Gly	Cys	Met	Ile	Leu	
225					230					235					240	
gcc	gat	ccc	acc	tac	ccg	ctc	acc	cct	gat	ccc	cgg	gac	gtg	tgg	gcg	768
Ala	Asp	Pro	Thr	Tyr	Pro	Leu	Thr	Pro	Asp	Pro	Arg	Asp	Val	Trp	Ala	
				245					250					255		
gcc	aag	cag	gca	gag	cgc	gcc	aac	tac	gcc	ttc	gga	gac	ctc	cac	gta	816
Ala	Lys	Gln	Ala	Glu	Arg	Ala	Asn	Tyr	Ala	Phe	Gly	Asp	Leu	His	Val	
			260				265						270			
cgt	ggt	gag	tac	ccc	gga	tac	ctg	cgg	cgg	acc	ctg	cgg	gac	aag	ggc	864
Arg	Gly	Glu	Tyr	Pro	Gly	Tyr	Leu	Arg	Arg	Thr	Leu	Arg	Asp	Lys	Gly	
		275					280					285				
atc	gag	ctg	gag	atc	acc	gag	gag	gac	cgc	gtg	ctg	ctg	cgg	gag	cac	912
Ile	Glu	Leu	Glu	Ile	Thr	Glu	Glu	Asp	Arg	Val	Leu	Leu	Arg	Glu	His	
	290					295					300					
acc	gtc	gac	ttc	gtc	tcc	ttc	tcc	tac	tac	atg	tnc	gtg	tgc	gag	acc	960
Thr	Val	Asp	Phe	Val	Ser	Phe	Ser	Tyr	Tyr	Met	Xaa	Val	Cys	Glu	Thr	
305					310					315					320	
gtc	acc	cag	tcg	gcc	gag	gcc	ggc	cgg	ggc	aac	ctc	atg	ggc	ggc	gtc	1008
Val	Thr	Gln	Ser	Ala	Glu	Ala	Gly	Arg	Gly	Asn	Leu	Met	Gly	Gly	Val	
				325			330							335		
ccc	aat	ccc	acc	ctc	gag	gcc	tcc	gag	tgg	gga	tgg	cag	atc	gac	ccg	1056
Pro	Asn	Pro	Thr	Leu	Glu	Ala	Ser	Glu	Trp	Gly	Trp	Gln	Ile	Asp	Pro	
			340				345						350			
gcg	ggc	ctg	cgc	acc	atc	ctg	aac	gac	tac	tgg	gac	cgc	tgg	ggc	aag	1104
Ala	Gly	Leu	Arg	Thr	Ile	Leu	Asn	Asp	Tyr	Trp	Asp	Arg	Trp	Gly	Lys	
		355				360						365				
cct	ctg	ttc	atc	gtc	gag	aac	ggc	ctg	gga	gcc	aag	gac	gtc	ctc	gtt	1152

Pro	Leu	Phe	Ile	Val	Glu	Asn	Gly	Leu	Gly	Ala	Lys	Asp	Val	Leu	Val		
370						375					380						
gac	gga	ccc	aac	ggt	ccc	acg	gtc	gag	gac	gac	tac	cgc	atc	gcc	tac	1200	
Asp	Gly	Pro	Asn	Gly	Pro	Thr	Val	Glu	Asp	Asp	Tyr	Arg	Ile	Ala	Tyr		
385					390					395					400		
atg	aac	gac	cac	ctg	gtc	cag	gtc	gcc	gag	gcc	att	gcc	gac	ggc	gtc	1248	
Met	Asn	Asp	His	Leu	Val	Gln	Val	Ala	Glu	Ala	Ile	Ala	Asp	Gly	Val		
				405					410					415			
gag	gtc	ctg	ggc	tac	acc	tcc	tgg	ggc	tgc	atc	gac	ctg	gtc	tcg	gcc	1296	
Glu	Val	Leu	Gly	Tyr	Thr	Ser	Trp	Gly	Cys	Ile	Asp	Leu	Val	Ser	Ala		
			420					425					430				
tcc	acc	gcc	cag	atg	tcc	aag	cgc	tac	ggg	ttc	atc	tac	gtg	gac	cgt	1344	
Ser	Thr	Ala	Gln	Met	Ser	Lys	Arg	Tyr	Gly	Phe	Ile	Tyr	Val	Asp	Arg		
		435					440					445					
gac	gac	ggc	ggc	aac	ggc	acc	ctg	gcc	cgc	tac	cgc	aag	aag	tcc	ttc	1392	
Asp	Asp	Gly	Gly	Asn	Gly	Thr	Leu	Ala	Arg	Tyr	Arg	Lys	Lys	Ser	Phe		
	450					455					460						
ggc	tgg	tac	cgc	gac	gtc	atc	gcc	tcc	aac	ggg	gcc	tcc	ctc	gtg	cct	1440	
Gly	Trp	Tyr	Arg	Asp	Val	Ile	Ala	Ser	Asn	Gly	Ala	Ser	Leu	Val	Pro		
465					470					475					480		
ccg	gtg	cag	gaa	ccg	ccg	cgg	ggg	tag								1467	
Pro	Val	Gln	Glu	Pro	Pro	Arg	Gly										
				485													

<210> 7  
 <211> 488  
 <212> PRT  
 <213> Actinomyces naeslundii  
  
 <220>  
 <221> misc\_feature  
 <222> (113)..(113)  
 <223> The 'Xaa' at location 113 stands for Leu.  
  
 <220>  
 <221> misc\_feature  
 <222> (148)..(148)  
 <223> The 'Xaa' at location 148 stands for Asp, Gly, Ala, or Val.  
  
 <220>  
 <221> misc\_feature  
 <222> (316)..(316)  
 <223> The 'Xaa' at location 316 stands for Tyr, Cys, Ser, or Phe.  
  
 <220>  
 <221> misc\_feature  
 <222> (339)..(339)  
 <223> nucleotide is uncertain  
  
 <220>  
 <221> misc\_feature  
 <222> (947)..(947)  
 <223> nucleotide is uncertain  
  
 <220>  
 <221> misc\_feature

[illegible]

Page 11

Ala Asp Pro Thr Tyr Pro Leu Thr Pro Asp Pro Arg Asp Val Trp Ala  
245 250 255

Ala Lys Gln Ala Glu Arg Ala Asn Tyr Ala Phe Gly Asp Leu His Val  
260 265 270

Arg Gly Glu Tyr Pro Gly Tyr Leu Arg Arg Thr Leu Arg Asp Lys Gly  
275 280 285

Ile Glu Leu Glu Ile Thr Glu Glu Asp Arg Val Leu Leu Arg Glu His  
290 295 300

Thr Val Asp Phe Val Ser Phe Ser Tyr Tyr Met Xaa Val Cys Glu Thr  
305 310 315 320

Val Thr Gln Ser Ala Glu Ala Gly Arg Gly Asn Leu Met Gly Gly Val  
325 330 335

Pro Asn Pro Thr Leu Glu Ala Ser Glu Trp Gly Trp Gln Ile Asp Pro  
340 345 350

Ala Gly Leu Arg Thr Ile Leu Asn Asp Tyr Trp Asp Arg Trp Gly Lys  
355 360 365

Pro Leu Phe Ile Val Glu Asn Gly Leu Gly Ala Lys Asp Val Leu Val  
370 375 380

Asp Gly Pro Asn Gly Pro Thr Val Glu Asp Asp Tyr Arg Ile Ala Tyr  
385 390 395 400

Met Asn Asp His Leu Val Gln Val Ala Glu Ala Ile Ala Asp Gly Val  
405 410 415

Glu Val Leu Gly Tyr Thr Ser Trp Gly Cys Ile Asp Leu Val Ser Ala  
420 425 430

Ser Thr Ala Gln Met Ser Lys Arg Tyr Gly Phe Ile Tyr Val Asp Arg  
435 440 445

Asp Asp Gly Gly Asn Gly Thr Leu Ala Arg Tyr Arg Lys Lys Ser Phe  
450 455 460

Gly Trp Tyr Arg Asp Val Ile Ala Ser Asn Gly Ala Ser Leu Val Pro  
465 470 475 480

Pro Val Gln Glu Pro Pro Arg Gly  
485

```
<210>      8
<211>    3072
<212>    DNA
<213>    Streptococcus salivarius

<220>
<221>    CDS
<222>    (392)..(2860)
<223>    1,6-alpha-glucanhydrolase
```

<400> 8																
aactgaggcc	gttgctccag	tagcgacaac	agaaataggt	ccatcaactg	ctactgttgc											60
gacagatact	gcaacaacag	cgacagcttc	tacaatcttt	tcacaagctg	tgccagcaga											120
aagtgctagc	tcagaaaacgc	ttgtagccag	tgaagcacta	gctcctgagt	cagctgctgt											180
ggaaaccatc	acatcatcat	ctgataatgc	tactgaagca	ggacgccatt	caactgctca											240
agtaacacca	gttacagaag	tgacagagca	aaacttgaat	ggtgatgcct	acttgacaga											300
tccagaaaaca	acaaaagcag	cttatagcaa	gacagatggt	gatattaatt	attccgttgt											360
tggtgtcta	at	g	atg	act	gtc	aac	ttg	aca	ctt							412
			Met	Thr	Val	Asn	Leu	Thr	Leu							
			1				5									
caa	cat	gct	tca	gaa	att	atc	ggt	caa	gat	aac	ggt	gac	ctt	acg	cta	460
Gln	His	Ala	Ser	Glu	Ile	Ile	Gly	Gln	Asp	Asn	Val	Asp	Leu	Thr	Leu	
		10					15					20				
gcg	gca	gga	gct	tca	gcc	aag	ggt	tca	aac	ttg	aca	gta	gcg	tca	gag	508
Ala	Ala	Gly	Ala	Ser	Ala	Lys	Val	Ser	Asn	Leu	Thr	Val	Ala	Ser	Glu	
		25				30					35					
tggttg	aca	aac	aat	aca	ggt	tac	ttg	gtg	aca	atc	agt	gtc	aac	gat		556
Trp	Leu	Thr	Asn	Asn	Thr	Gly	Tyr	Leu	Val	Thr	Ile	Ser	Val	Asn	Asp	
40				45					50					55		
aaa	tca	ggc	aat	gtc	ttg	tca	agc	aag	cgc	gct	ggc	ttg	tct	gtt	gaa	604
Lys	Ser	Gly	Asn	Val	Leu	Ser	Ser	Lys	Arg	Ala	Gly	Leu	Ser	Val	Glu	
			60					65						70		
gat	gat	tggtga	aca	ggt	ttc	cca	cgt	tac	ggt	atc	gta	gca	gggt	tca	cca	652
Asp	Asp	Trp	Thr	Val	Phe	Pro	Arg	Tyr	Gly	Ile	Val	Ala	Gly	Ser	Pro	
			75				80						85			
act	gat	caa	aac	agt	att	ctt	ggt	aaa	aat	ctt	gaa	gcc	tac	cgt	aaa	700
Thr	Asp	Gln	Asn	Ser	Ile	Leu	Val	Lys	Asn	Leu	Glu	Ala	Tyr	Arg	Lys	
		90					95					100				
gag	ctt	gag	ctc	atg	aag	tct	atg	aat	atc	aac	tca	tat	ttc	ttc	tat	748
Glu	Leu	Glu	Leu	Met	Lys	Ser	Met	Asn	Ile	Asn	Ser	Tyr	Phe	Phe	Tyr	
		105				110					115					
gat	gct	tat	aat	gaa	gct	aca	gat	cct	ttc	cca	gaa	gggt	gtc	gat	agc	796
Asp	Ala	Tyr	Asn	Glu	Ala	Thr	Asp	Pro	Phe	Pro	Glu	Gly	Val	Asp	Ser	
120				125				130							135	
ttt	gtt	caa	aaa	tggtga	aat	acc	tggtga	agt	cac	act	cag	gtt	gac	act	aag	844
Phe	Val	Gln	Lys	Trp	Asn	Thr	Trp	Ser	His	Thr	Gln	Val	Asp	Thr	Lys	
				140				145					150			

gct Ala	ggt Val	aaa Lys	gaa Glu 155	ttg Leu	ggt Val	gat Asp	caa Gln	ggt Val 160	cat His	aag Lys	tca Ser	ggg Gly	gct Ala 165	ggt Val	gcc Ala	892
atg Met	ctt Leu	tat Tyr 170	aac Asn	atg Met	att Ile	tca Ser	gca Ala 175	gat Asp	tca Ser	aat Asn	cca Pro	aag Lys 180	aat Asn	ccg Pro	gcc Ala	940
ctt Leu	cca Pro 185	ctt Leu	gct Ala	gct Ala	ttg Leu	gct Ala 190	tat Tyr	aac Asn	ttc Phe	tac Tyr	gat Asp 195	agc Ser	ttt Phe	ggt Gly	aag Lys	988
aag Lys 200	ggg Gly	gaa Glu	ccg Pro	atg Met	act Thr 205	tac Tyr	act Thr	atc Ile	ggg Gly	gat Asp 210	aac Asn	cca Pro	act Thr	caa Gln	ggt Val 215	1036
tac Tyr	tat Tyr	gat Asp	ccg Pro	gcg Ala 220	aat Asn	cca Pro	gat Asp	tgg Trp	caa Gln 225	aaa Lys	tac Tyr	atc Ile	gca Ala	ggg Gly 230	gtc Val	1084
atg Met	aaa Lys	tca Ser	gct Ala 235	atg Met	gat Asp	cgt Arg	atg Met	gga Gly 240	ttc Phe	gat Asp	ggg Gly	tgg Trp	caa Gln 245	ggg Gly	gat Asp	1132
aca Thr	att Ile	ggg Gly 250	gac Asp	aac Asn	cgt Arg	gtg Val	act Thr 255	gat Asp	tat Tyr	gag Glu	cac His	cgt Arg 260	aac Asn	agc Ser	aca Thr	1180
gac Asp	gag Glu 265	gct Ala	gac Asp	tca Ser	cac His	atg Met 270	atg Met	tct Ser	gat Asp	tca Ser	tat Tyr 275	gcg Ala	tca Ser	ttt Phe	att Ile	1228
aat Asn 280	gcc Ala	atg Met	aag Lys	gac Asp	ctc Leu 285	atc Ile	ggg Gly	gaa Glu	aag Lys	tac Tyr 290	tac Tyr	atc Ile	aca Thr	atc Ile	aat Asn 295	1276
gat Asp	ggt Val	aat Asn	ggg Gly 300	ggg Gly	aat Asn	gat Asp	gat Asp	aaa Lys	cta Leu 305	gcc Ala	aag Lys	gca Ala	cgt Arg	caa Gln 310	gat Asp	1324
ggt Val	ggt Val	tat Tyr	aat Asn 315	gag Glu	ctt Leu	tgg Trp	aca Thr	aac Asn 320	ggg Gly	ggg Gly	tca Ser	ggt Val	att Ile 325	cca Pro	gga Gly	1372
cgt Arg	atg Met	cag Gln 330	ggt Val	gcc Ala	tat Tyr	ggg Gly	gat Asp 335	ttg Leu	aaa Lys	gca Ala	cgt Arg	atc Ile 340	gat Asp	atg Met	gta Val	1420
cgc Arg	aat Asn 345	aaa Lys	act Thr	ggg Gly	aaa Lys	tca Ser 350	ctt Leu	atc Ile	ggt Val	ggg Gly	gcc Ala 355	tac Tyr	atg Met	gaa Glu	gaa Glu	1468
cca Pro 360	ggg Gly	att Ile	gat Asp	tat Tyr	act Thr 365	ggt Val	cct Pro	ggc Gly	gga Gly	aaa Lys 370	gca Ala	act Thr	aac Asn	ggg Gly	gct Ala 375	1516
ggg Gly	aaa Lys	gat Asp	gcc Ala	ctt Leu 380	gct Ala	ggg Gly	aaa Lys	cca Pro	ttg Leu 385	caa Gln	gct Ala	gat Asp	gcg Ala	act Thr 390	ctt Leu	1564
ctc Leu	gta Val	gat Asp	gcg Ala 395	aca Thr	gta Val	gct Ala	gca Ala	gca Ala 400	ggg Gly	ggg Gly	tat Tyr	cac His	atg Met 405	tcc Ser	att Ile	1612

gca Ala	gcc Ala	ctt Leu 410	gca Ala	aat Asn	gct Ala	aat Asn	gcg Ala 415	gcc Ala	ctt Leu	aac Asn	gtc Val	ctt Leu 420	caa Gln	agt Ser	gcc Ala	1660
tat Tyr	tac Tyr 425	cca Pro	acg Thr	caa Gln	tac Tyr	ctc Leu 430	agt Ser	gtg Val	gct Ala	aaa Lys	gac Asp 435	act Thr	att Ile	cgt Arg	aag Lys	1708
ctt Leu 440	tac Tyr	aat Asn	tac Tyr	caa Gln	cag Gln 445	ttt Phe	atc Ile	act Thr	gct Ala	tat Tyr 450	gaa Glu	aat Asn	ctt Leu	ctc Leu	cgc Arg 455	1756
ggc Gly	gag Glu	ggc Gly	gtg Val	aca Thr 460	aac Asn	agc Ser	act Thr	cag Gln	gct Ala 465	gta Val	tct Ser	aca Thr	aag Lys	aat Asn	gct Ala 470	1804
tct Ser	ggc Gly	gaa Glu	atc Ile 475	ctt Leu	tct Ser	aaa Lys	gat Asp	gct Ala 480	ctt Leu	ggc Gly	gtg Val	aca Thr	gga Gly 485	gat Asp	caa Gln	1852
gtt Val	tgg Trp	aca Thr 490	ttt Phe	gct Ala	aaa Lys	tca Ser	gga Gly 495	aaa Lys	ggc Gly	ttc Phe	tca Ser	act Thr 500	gtt Val	caa Gln	atg Met	1900
att Ile	aat Asn 505	atg Met	atg Met	ggc Gly	atc Ile	aat Asn 510	gcg Ala	ggc Gly	tgg Trp	cat His	aat Asn 515	gaa Glu	gag Glu	ggc Gly	tat Tyr	1948
gcg Ala 520	gac Asp	aat Asn	aaa Lys	aca Thr	ccg Pro 525	gac Asp	gca Ala	caa Gln	gaa Glu	aat Asn 530	ctc Leu	aca Thr	gtt Val	cgt Arg	ctt Leu 535	1996
agc Ser	cta Leu	gca Ala	ggc Gly	aaa Lys 540	aca Thr	gcc Ala	caa Gln	gaa Glu	gca Ala 545	gct Ala	aaa Lys	att Ile	gct Ala	gat Asp 550	caa Gln	2044
gtc Val	tat Tyr	gtg Val	acg Thr 555	tca Ser	ccg Pro	gat Asp	gat Asp	tgg Trp 560	gca Ala	act Thr	tca Ser	agc Ser	atg Met 565	aag Lys	aag Lys	2092
gca Ala	caa Gln	gca Ala 570	agc Ser	ctt Leu	gaa Glu	aca Thr	gat Asp 575	gaa Glu	aat Asn	ggc Gly	caa Gln	cca Pro 580	gtg Val	ctt Leu	gtc Val	2140
att Ile	tca Ser 585	gtt Val	cct Pro	aaa Lys	cta Leu	act Thr 590	ctt Leu	tgg Trp	aac Asn	atg Met	ctt Leu 595	tat Tyr	atc Ile	aag Lys	gaa Glu	2188
gac Asp 600	aca Thr	aca Thr	gca Ala	aca Thr	ccg Pro 605	gta Val	gaa Glu	cca Pro	gtt Val	act Thr 610	aac Asn	caa Gln	gct Ala	ggc Gly	aag Lys 615	2236
aaa Lys	gta Val	gat Asp	aat Asn	acc Thr 620	gta Val	aca Thr	tct Ser	gaa Glu	gca Ala 625	agc Ser	tca Ser	gaa Glu	aca Thr	gct Ala 630	aaa Lys	2284
tca Ser	gaa Glu	aat Asn	aca Thr 635	aca Thr	gta Val	aat Asn	aaa Lys	ggc Gly 640	tca Ser	gag Glu	gct Ala	cca Pro	act Thr 645	gat Asp	acg Thr	2332
aaa Lys	cca Pro	tct Ser 650	gtt Val	gaa Glu	gct Ala	cct Pro	aaa Lys 655	cta Leu	gat Asp	gaa Glu	aca Thr	act Thr 660	aaa Lys	cca Pro	gca Ala	2380

cca tca gtt gac gag tta gta aac tca gca gct gtt cca gtg gcg ata	2428
Pro Ser Val Asp Glu Leu Val Asn Ser Ala Ala Val Pro Val Ala Ile	
665 670 675	
gct gtg tca gag acc gca cat gat aag aaa gat gac aac tca gta tct	2476
Ala Val Ser Glu Thr Ala His Asp Lys Lys Asp Asp Asn Ser Val Ser	
680 685 690 695	
aat acg gat caa ggt aca gta gca tca gat tca atc act aca cca gct	2524
Asn Thr Asp Gln Gly Thr Val Ala Ser Asp Ser Ile Thr Thr Pro Ala	
700 705 710	
tca gag gct gca agc aca gct gcc tca aca gtc tca tca gaa gta tca	2572
Ser Glu Ala Ala Ser Thr Ala Ala Ser Thr Val Ser Ser Glu Val Ser	
715 720 725	
gaa agt gta aca gta tca tcg gaa cca tca gaa act gaa aat agt tca	2620
Glu Ser Val Thr Val Ser Ser Glu Pro Ser Glu Thr Glu Asn Ser Ser	
730 735 740	
gaa gca tca act tca gag tca gca act cca acg acg aca gca att tca	2668
Glu Ala Ser Thr Ser Glu Ser Ala Thr Pro Thr Thr Thr Ala Ile Ser	
745 750 755	
gaa tca cat gca gta gtt gaa cca gtg gct tct ttg aca gaa tca gag	2716
Glu Ser His Ala Val Val Glu Pro Val Ala Ser Leu Thr Glu Ser Glu	
760 765 770 775	
agt cag gca agc act agc ctt gtt tca gaa act aca agc aca att gtc	2764
Ser Gln Ala Ser Thr Ser Leu Val Ser Glu Thr Thr Ser Thr Ile Val	
780 785 790	
tca gtt gct ccg tca gaa gta tca gaa agc aca tca gag gaa gtc atc	2812
Ser Val Ala Pro Ser Glu Val Ser Glu Ser Thr Ser Glu Glu Val Ile	
795 800 805	
ctt atg gac tat cag aaa aca tca ata gtt gga ata gac tct ctg tag	2860
Leu Met Asp Tyr Gln Lys Thr Ser Ile Val Gly Ile Asp Ser Leu	
810 815 820	
ctcctcgcgt ctcagaaacc ttaccaagta cttctgaaac gattacagaa gcagcatcac	2920
tcttttagcaa ctatgcaaga tattcagaaa cagcaagctc agaatctcac tctatggtag	2980
cagcttcttc agaagtttct attgaaaaat tagcagtatc tatcttgaaa gatactgagg	3040
gaggcttgta tgatgcaaca acaatcagaa at	3072

<210> 9  
 <211> 822  
 <212> PRT  
 <213> Streptococcus salivarius

<400> 9

Met Thr Val Asn Leu Thr Leu Gln His Ala Ser Glu Ile Ile Gly Gln
1 5 10 15

Asp Asn Val Asp Leu Thr Leu Ala Ala Gly Ala Ser Ala Lys Val Ser
20 25 30



Asn Leu Thr Val Ala Ser Glu Trp Leu Thr Asn Asn Thr Gly Tyr Leu  
 35 40 45  
 Val Thr Ile Ser Val Asn Asp Lys Ser Gly Asn Val Leu Ser Ser Lys  
 50 55 60  
 Arg Ala Gly Leu Ser Val Glu Asp Asp Trp Thr Val Phe Pro Arg Tyr  
 65 70 75 80  
 Gly Ile Val Ala Gly Ser Pro Thr Asp Gln Asn Ser Ile Leu Val Lys  
 85 90 95  
 Asn Leu Glu Ala Tyr Arg Lys Glu Leu Glu Leu Met Lys Ser Met Asn  
 100 105 110  
 Ile Asn Ser Tyr Phe Phe Tyr Asp Ala Tyr Asn Glu Ala Thr Asp Pro  
 115 120 125  
 Phe Pro Glu Gly Val Asp Ser Phe Val Gln Lys Trp Asn Thr Trp Ser  
 130 135 140  
 His Thr Gln Val Asp Thr Lys Ala Val Lys Glu Leu Val Asp Gln Val  
 145 150 155 160  
 His Lys Ser Gly Ala Val Ala Met Leu Tyr Asn Met Ile Ser Ala Asp  
 165 170 175  
 Ser Asn Pro Lys Asn Pro Ala Leu Pro Leu Ala Ala Leu Ala Tyr Asn  
 180 185 190  
 Phe Tyr Asp Ser Phe Gly Lys Lys Gly Glu Pro Met Thr Tyr Thr Ile  
 195 200 205  
 Gly Asp Asn Pro Thr Gln Val Tyr Tyr Asp Pro Ala Asn Pro Asp Trp  
 210 215 220  
 Gln Lys Tyr Ile Ala Gly Val Met Lys Ser Ala Met Asp Arg Met Gly  
 225 230 235 240  
 Phe Asp Gly Trp Gln Gly Asp Thr Ile Gly Asp Asn Arg Val Thr Asp  
 245 250 255  
 Tyr Glu His Arg Asn Ser Thr Asp Glu Ala Asp Ser His Met Met Ser  
 260 265 270  
 Asp Ser Tyr Ala Ser Phe Ile Asn Ala Met Lys Asp Leu Ile Gly Glu  
 275 280 285

Lys Tyr Tyr Ile Thr Ile Asn Asp Val Asn Gly Gly Asn Asp Asp Lys  
290 295 300

Leu Ala Lys Ala Arg Gln Asp Val Val Tyr Asn Glu Leu Trp Thr Asn  
305 310 315 320

Gly Gly Ser Val Ile Pro Gly Arg Met Gln Val Ala Tyr Gly Asp Leu  
325 330 335

Lys Ala Arg Ile Asp Met Val Arg Asn Lys Thr Gly Lys Ser Leu Ile  
340 345 350

Val Gly Ala Tyr Met Glu Glu Pro Gly Ile Asp Tyr Thr Val Pro Gly  
355 360 365

Gly Lys Ala Thr Asn Gly Ala Gly Lys Asp Ala Leu Ala Gly Lys Pro  
370 375 380

Leu Gln Ala Asp Ala Thr Leu Leu Val Asp Ala Thr Val Ala Ala Ala  
385 390 395 400

Gly Gly Tyr His Met Ser Ile Ala Ala Leu Ala Asn Ala Asn Ala Ala  
405 410 415

Leu Asn Val Leu Gln Ser Ala Tyr Tyr Pro Thr Gln Tyr Leu Ser Val  
420 425 430

Ala Lys Asp Thr Ile Arg Lys Leu Tyr Asn Tyr Gln Gln Phe Ile Thr  
435 440 445

Ala Tyr Glu Asn Leu Leu Arg Gly Glu Gly Val Thr Asn Ser Thr Gln  
450 455 460

Ala Val Ser Thr Lys Asn Ala Ser Gly Glu Ile Leu Ser Lys Asp Ala  
465 470 475 480

Leu Gly Val Thr Gly Asp Gln Val Trp Thr Phe Ala Lys Ser Gly Lys  
485 490 495

Gly Phe Ser Thr Val Gln Met Ile Asn Met Met Gly Ile Asn Ala Gly  
500 505 510

Trp His Asn Glu Glu Gly Tyr Ala Asp Asn Lys Thr Pro Asp Ala Gln  
515 520 525

Glu Asn Leu Thr Val Arg Leu Ser Leu Ala Gly Lys Thr Ala Gln Glu  
530 535 540

Ala Ala Lys Ile Ala Asp Gln Val Tyr Val Thr Ser Pro Asp Asp Trp  
545 550 555 560

Ala Thr Ser Ser Met Lys Lys Ala Gln Ala Ser Leu Glu Thr Asp Glu  
565 570 575

Asn Gly Gln Pro Val Leu Val Ile Ser Val Pro Lys Leu Thr Leu Trp  
580 585 590

Asn Met Leu Tyr Ile Lys Glu Asp Thr Thr Ala Thr Pro Val Glu Pro  
595 600 605

Val Thr Asn Gln Ala Gly Lys Lys Val Asp Asn Thr Val Thr Ser Glu  
610 615 620

Ala Ser Ser Glu Thr Ala Lys Ser Glu Asn Thr Thr Val Asn Lys Gly  
625 630 635 640

Ser Glu Ala Pro Thr Asp Thr Lys Pro Ser Val Glu Ala Pro Lys Leu  
645 650 655

Asp Glu Thr Thr Lys Pro Ala Pro Ser Val Asp Glu Leu Val Asn Ser  
660 665 670

Ala Ala Val Pro Val Ala Ile Ala Val Ser Glu Thr Ala His Asp Lys  
675 680 685

Lys Asp Asp Asn Ser Val Ser Asn Thr Asp Gln Gly Thr Val Ala Ser  
690 695 700

Asp Ser Ile Thr Thr Pro Ala Ser Glu Ala Ala Ser Thr Ala Ala Ser  
705 710 715 720

Thr Val Ser Ser Glu Val Ser Glu Ser Val Thr Val Ser Ser Glu Pro  
725 730 735

Ser Glu Thr Glu Asn Ser Ser Glu Ala Ser Thr Ser Glu Ser Ala Thr  
740 745 750

Pro Thr Thr Thr Ala Ile Ser Glu Ser His Ala Val Val Glu Pro Val  
755 760 765

Ala Ser Leu Thr Glu Ser Glu Ser Gln Ala Ser Thr Ser Leu Val Ser  
770 775 780

Glu Thr Thr Ser Thr Ile Val Ser Val Ala Pro Ser Glu Val Ser Glu  
785 790 795 800

Ser Thr Ser Glu Glu Val Ile Leu Met Asp Tyr Gln Lys Thr Ser Ile  
805 810 815

Val Gly Ile Asp Ser Leu  
820

<210> 10  
<211> 2220  
<212> DNA  
<213> *Trichoderma longibrachiatum*

<400> 10

aaggttagcc aagaacaata gccgataaag atagcctcat taaacggaat gagctagtag 60  
gcaaagtcag cgaatgtgta tatataaagg ttcgaggtec gtgcctccct catgctctcc 120  
ccatctactc atcaactcag atcctccagg agacttgtag accatctttt gaggcacaga 180  
aacccaatag tcaaccgcgg actggcatca tgtatcggaa gttggccgtc atcacggcct 240  
tcttgccac agctcgtgct cagtcggcct gcactctcca atcggagact caccgcctc 300  
tgacatggca gaaatgctcg tctggtggca cttgcactca acagacaggc tccgtggtca 360  
tcgacgcaa ctggcgctgg actcacgcta cgaacagcag cacgaactgc tacgatggca 420  
acacttgag ctcgacccta tgtcctgaca acgagacctg cggaagaac tgctgtctgg 480  
acggtgccgc ctacgcgtcc acgtacggag ttaccacgag cggtaacagc ctctccattg 540  
gctttgtcac ccagtctgcg cagaagaacg ttggcgctcg cctttacctt atggcgagcg 600  
acacgaccta ccaggaattc accctgcttg gcaacgagtt ctctttcgat gttgatgttt 660  
cgcagctgcc gtaagtgact taccatgaac ccctgacgta tcttcttggt ggctccagc 720  
tgactggcca atttaaggtg cggcttgaac ggagctctct acttcgtgtc catggacgcg 780  
gatggtggcg tgagcaagta tcccaccaac aacgctggcg ccaagtacgg cacggggtag 840  
tgtgacagcc agtgtccccg cgatctgaag ttcatcaatg gccaggcaa cgttgagggc 900  
tgaggagccgt catccaacaa cgcaaacacg ggcattggag gacacggaag ctgctgctct 960  
gagatggata tctgggaggc caactccatc tccgaggctc ttacccccca cccttgcaag 1020  
actgtcggcc aggagatctg cgagggtgat ggtgcgggc gaacttactc cgataacaga 1080  
tatggcgga cttgcgatcc cgatggctgc gactggaacc cataccgctt gggcaacacc 1140  
agcttctacg gccctggctc aagctttacc ctcgatacca ccaagaaatt gaccgttgct 1200  
accagttcg agacgtcggg tgccatcaac cgatactatg tccagaatgg cgtcactttc 1260  
cagcagccca acgccgagct tggtagttac tctggcaacg agctcaacga tgattactgc 1320  
acagctgagg agacagaatt cggcggatct ctttctcaga caagggcggc ctgactcagt 1380  
tcaagaaggc tacctctggc ggcattggtc tggatcatgag tctgtgggat gatgtgagtt 1440





Cys Ala Ser Thr Pro Phe Pro Thr Leu Ile Thr Ala Pro Gly Pro Ser  
 345 350 355  
 gcg tcc gtc gct cgc atc ccc ccg ccg ccg tcc ccc aac taa 1152  
 Ala Ser Val Ala Arg Ile Pro Pro Pro Pro Ser Pro Asn  
 360 365 370  
 gctatgtcta tgctggacat gctctcggtt ctacctcgtc ggtatcgctc cacggttatc 1212  
 tcgcgttttgc atcatgtata cctgctcgtg gaatatacaa agtgggtctat c 1263

<210> 12  
 <211> 372  
 <212> PRT  
 <213> Phanerochaete chrysosporium

<400> 12

Met Ala Phe Lys Gln Leu Leu Ala Ala Leu Ser Val Ala Leu Thr Leu  
 1 5 10 15  
 Gln Val Thr Gln Ala Ala Pro Asn Leu Asp Lys Arg Val Ala Cys Pro  
 20 25 30  
 Asp Gly Val His Thr Ala Ser Asn Ala Ala Cys Cys Ala Trp Phe Pro  
 35 40 45  
 Val Leu Asp Asp Ile Gln Gln Asn Leu Phe His Gly Gly Gln Cys Gly  
 50 55 60  
 Ala Glu Ala His Glu Ala Leu Arg Met Val Phe His Asp Ser Ile Ala  
 65 70 75 80  
 Ile Ser Pro Lys Leu Gln Ser Gln Gly Lys Phe Gly Gly Gly Gly Ala  
 85 90 95  
 Asp Gly Ser Ile Ile Thr Phe Ser Ser Ile Glu Thr Thr Tyr His Pro  
 100 105 110  
 Asn Ile Gly Leu Asp Glu Val Val Ala Ile Gln Lys Pro Phe Ile Ala  
 115 120 125  
 Lys His Gly Val Thr Arg Gly Asp Phe Ile Ala Phe Ala Gly Ala Val  
 130 135 140  
 Gly Val Ser Asn Cys Pro Gly Ala Pro Gln Met Gln Phe Phe Leu Gly  
 145 150 155 160  
 Arg Pro Glu Ala Thr Gln Ala Ala Pro Asp Gly Leu Val Pro Glu Pro  
 165 170 175  
 Phe His Thr Ile Asp Gln Val Leu Ala Arg Met Leu Asp Ala Gly Gly

180

185

190

Phe Asp Glu Ile Glu Thr Val Trp Leu Leu Ser Ala His Ser Ile Ala  
195 200 205

Ala Ala Asn Asp Val Asp Pro Thr Ile Ser Gly Leu Pro Phe Asp Ser  
210 215 220

Thr Pro Gly Gln Phe Asp Ser Gln Phe Phe Val Glu Thr Gln Leu Arg  
225 230 235 240

Gly Thr Ala Phe Pro Gly Lys Thr Gly Ile Gln Gly Thr Val Met Ser  
245 250 255

Pro Leu Lys Gly Glu Met Arg Leu Gln Thr Asp His Leu Phe Ala Arg  
260 265 270

Asp Ser Arg Thr Ala Cys Glu Trp Gln Ser Phe Val Asn Asn Gln Thr  
275 280 285

Lys Leu Gln Glu Asp Phe Gln Phe Ile Phe Thr Ala Leu Ser Thr Leu  
290 295 300

Gly His Asp Met Asn Ala Met Ile Asp Cys Ser Glu Val Ile Pro Ala  
305 310 315 320

Pro Lys Pro Val Asn Phe Gly Pro Ser Phe Phe Pro Ala Gly Lys Thr  
325 330 335

His Ala Asp Ile Glu Gln Ala Cys Ala Ser Thr Pro Phe Pro Thr Leu  
340 345 350

Ile Thr Ala Pro Gly Pro Ser Ala Ser Val Ala Arg Ile Pro Pro Pro  
355 360 365

Pro Ser Pro Asn  
370

<210> 13  
<211> 1285  
<212> DNA  
<213> Phanerochaete chrysosporium

<220>  
<221> CDS  
<222> (34)..(1149) <223> CKG5 ligninase precursor  
<400> 13

gtcagactct ccaacggttg cctttggaca gac atg gcc ttc aag aag ctc ctt  
Met Ala Phe Lys Lys Leu Leu

54



5

[illegible]

gac ttc ctg atc gct cgt gac gcg cgc acc gcc tgc gag tgg cag tcg 870  
 Asp Phe Leu Ile Ala Arg Asp Ala Arg Thr Ala Cys Glu Trp Gln Ser  
 265 270 275

ttc gtc aac aac cag tcc aag ctc gtc tcc gac ttc caa ttc atc ttc 918  
 Phe Val Asn Asn Gln Ser Lys Leu Val Ser Asp Phe Gln Phe Ile Phe  
 280 285 290 295

ctc gcc ctc act cag ctc ggc cag gac ccg gat gcg atg acc gac tgc 966  
 Leu Ala Leu Thr Gln Leu Gly Gln Asp Pro Asp Ala Met Thr Asp Cys  
 300 305 310

tct gct gtc atc ccc atc tcc aag ccc gcc ccg aac aac acc ccc gga 1014  
 Ser Ala Val Ile Pro Ile Ser Lys Pro Ala Pro Asn Asn Thr Pro Gly  
 315 320 325

ttc tcc ttc ttc ccg ccc ggc atg acg atg gac gat gtc gag cag gct 1062  
 Phe Ser Phe Phe Pro Pro Gly Met Thr Met Asp Asp Val Glu Gln Ala  
 330 335 340

tgc gcc gag acg ccc ttc ccg act ctc tcg act ctc cct ggc ccc gcg 1110  
 Cys Ala Glu Thr Pro Phe Pro Thr Leu Ser Thr Leu Pro Gly Pro Ala  
 345 350 355

acc tcc gtc gct cgc atc cct cct cct cct ggt gct taa gcagccatca 1159  
 Thr Ser Val Ala Arg Ile Pro Pro Pro Pro Gly Ala  
 360 365 370

gacttcggat cacaccccggt tattggcaac ggaaatttag aacgaagatc gtccagtgtt 1219

ttgaagtaga aatgtgcttg tactgtgtaa acagctcttt tgacgaaata cactctgatt 1279

tcgtcg 1285

<210> 14  
 <211> 371  
 <212> PRT  
 <213> Phanerochaete chrysosporium

<400> 14

Met Ala Phe Lys Lys Leu Leu Ala Val Leu Thr Ala Ala Leu Ser Leu  
 1 5 10 15

Arg Ala Ala Gln Gly Ala Ala Val Glu Lys Arg Ala Thr Cys Ser Asn  
 20 25 30

Gly Lys Val Val Pro Ala Ala Ser Cys Cys Thr Trp Phe Asn Val Leu  
 35 40 45

Ser Asp Ile Gln Glu Asn Leu Phe Asn Gly Gly Gln Cys Gly Ala Glu  
 50 55 60

Ala His Glu Ser Ile Arg Leu Val Phe His Asp Ala Ile Ala Ile Ser  
 65 70 75 80

Pro Ala Met Glu Pro Gln Ala Ser Ser Val Arg Gly Ala Asp Gly Ser  
85 90 95

Ile Met Ile Phe Asp Glu Ile Glu Thr Asn Phe His Pro Asn Ile Gly  
100 105 110

Leu Asp Glu Ile Val Arg Leu Gln Lys Pro Phe Val Gln Lys His Gly  
115 120 125

Val Thr Pro Gly Asp Phe Ile Ala Phe Ala Gly Ala Val Ala Leu Ser  
130 135 140

Asn Cys Pro Gly Ala Pro Gln Met Asn Phe Phe Thr Gly Arg Ala Pro  
145 150 155 160

Ala Thr Gln Pro Ala Pro Asp Gly Leu Val Pro Glu Pro Phe His Ser  
165 170 175

Val Asp Gln Ile Ile Asp Arg Val Phe Asp Ala Gly Glu Phe Asp Glu  
180 185 190

Leu Glu Leu Val Trp Met Leu Ser Ala His Ser Val Ala Ala Ala Asn  
195 200 205

Asp Ile Asp Pro Asn Ile Gln Gly Leu Pro Phe Asp Ser Thr Pro Gly  
210 215 220

Ile Phe Asp Ser Gln Phe Phe Val Glu Thr Gln Leu Ala Gly Thr Gly  
225 230 235 240

Phe Thr Gly Gly Ser Asn Asn Gln Gly Glu Val Ser Ser Pro Leu Pro  
245 250 255

Gly Glu Met Arg Leu Gln Ser Asp Phe Leu Ile Ala Arg Asp Ala Arg  
260 265 270

Thr Ala Cys Glu Trp Gln Ser Phe Val Asn Asn Gln Ser Lys Leu Val  
275 280 285

Ser Asp Phe Gln Phe Ile Phe Leu Ala Leu Thr Gln Leu Gly Gln Asp  
290 295 300

Pro Asp Ala Met Thr Asp Cys Ser Ala Val Ile Pro Ile Ser Lys Pro  
305 310 315 320

Ala Pro Asn Asn Thr Pro Gly Phe Ser Phe Phe Pro Pro Gly Met Thr  
325 330 335

Met Asp Asp Val Glu Gln Ala Cys Ala Glu Thr Pro Phe Pro Thr Leu  
 340 345 350

Ser Thr Leu Pro Gly Pro Ala Thr Ser Val Ala Arg Ile Pro Pro Pro  
 355 360 365

Pro Gly Ala  
 370

<210> 15  
 <211> 360  
 <212> DNA  
 <213> Solanum tuberosum

<400> 15

tgaccctaga cttgtccatc ttctggattg gccaaagttaa ttaatgtatg aaataaaaagg 60  
 atgcacacat agtgacatgc taatcactat aatgtgggca tcaaagttgt gtgttatgtg 120  
 taataactaa ttatctgaat aagagaaaga gagatcatcc atatttctta tcctaaatga 180  
 atgacagtgt ctttataatt ctttgatgaa cagatgcatt ttattaacca attccatata 240  
 catataaata ttaatcatat ataattaata tcaattgggt agcaaaaacc aaatctagtc 300  
 taggtgtgtt ttgctaatta tgggggatag agcaaaaaag aaactaacgt ctcaagaatc 360

<210> 16  
 <211> 2521  
 <212> DNA  
 <213> Agrobacterium tumefaciens

<220>  
 <221> CDS  
 <222> (585)..(1826)  
 <223> nopaline synthetase

<400> 16

tagccgaccc agacgagcca agggatcttt ttggaatgct gctccgctgt caggctttcc 60  
 gacgtttggg tggttgaaca gaagtcatta tcgtacggaa tgccaagcac tcccgagggg 120  
 aaccctgtgg ttggcatgca catacaaatg gacgaacgga taaacctttt cagccctttt 180  
 taaatatccg ttattctaata aaacgctctt ttctcttagg tttaccgccc aatataatcct 240  
 gtcaaacact gatagtttaa actgaaggcg ggaaacgaca atctgatcat gagcggagaa 300  
 ttaagggagt cacgttatga cccccgccga tgacgcggga caagccgttt tacgtttgga 360  
 actgacagaa ccgcaacgat tgaaggagcc actcagccgc gggtttcttg agtttaatga 420  
 gctaagcaca tacgtcagaa accattattg cgcgttcaaa agtcgcctaa ggtcactatc 480  
 agctagcaaa tatttcttgt caaaaatgct ccaactgacgt tccataaatt cccctcggtg 540  
 tccaattaga gtctcatatt cactctcaat ccaaataatc tgca atg gca att acc 596  
 Met Ala Ile Thr

tta tcc gca act tct tta cct att tcc gcc gca gat cac cat ccg ctt	644
Leu Ser Ala Thr Ser Leu Pro Ile Ser Ala Ala Asp His His Pro Leu	
5 10 15 20	
ccc ttg acc gta ggt gtc ctc ggt tct ggt cac gcg ggg act gca tta	692
Pro Leu Thr Val Gly Val Leu Gly Ser Gly His Ala Gly Thr Ala Leu	
25 30 35	
gcg gct tgg ttc gcc tcc cgg cat gtt ccc acg gcg ctg tgg gca cca	740
Ala Ala Trp Phe Ala Ser Arg His Val Pro Thr Ala Leu Trp Ala Pro	
40 45 50	
gca gat cat cca gga tcg atc tca gca atc aag gcc aat gaa gga gtt	788
Ala Asp His Pro Gly Ser Ile Ser Ala Ile Lys Ala Asn Glu Gly Val	
55 60 65	
atc acc acc gag gga atg att aac ggt cca ttt agg gtc tca gcc tgt	836
Ile Thr Thr Glu Gly Met Ile Asn Gly Pro Phe Arg Val Ser Ala Cys	
70 75 80	
gat gac ctt gcc gca gtt att cgc tcc agc cgt gta ctg att att gta	884
Asp Asp Leu Ala Ala Val Ile Arg Ser Ser Arg Val Leu Ile Ile Val	
85 90 95 100	
acc cgt gcg gac gtt cac gac agc ttc gtc aac gaa ctc gcc aac ttc	932
Thr Arg Ala Asp Val His Asp Ser Phe Val Asn Glu Leu Ala Asn Phe	
105 110 115	
aac ggc gaa ctc gca aca aag gat att gtc gtc gtg tgc ggc cat ggc	980
Asn Gly Glu Leu Ala Thr Lys Asp Ile Val Val Val Cys Gly His Gly	
120 125 130	
ttc tcc atc aag tac gag aga cag ctg cga ttc aag cga ata ttc gag	1028
Phe Ser Ile Lys Tyr Glu Arg Gln Leu Arg Phe Lys Arg Ile Phe Glu	
135 140 145	
acg gat aat tcg ccc ata acg tct aag cta tcg gat caa aaa aaa tgt	1076
Thr Asp Asn Ser Pro Ile Thr Ser Lys Leu Ser Asp Gln Lys Lys Cys	
150 155 160	
aac gtc aac atc aag gaa atg aaa gcg tct ttc gga ctg tca tgt ttc	1124
Asn Val Asn Ile Lys Glu Met Lys Ala Ser Phe Gly Leu Ser Cys Phe	
165 170 175 180	
cca att cat cgc gat gat gct ggc gtg att gat cta ccc gaa gat acc	1172
Pro Ile His Arg Asp Asp Ala Gly Val Ile Asp Leu Pro Glu Asp Thr	
185 190 195	
aag aac atc ttt gcc cag cta ttt tcc gct aga atc atc tgc atc ccg	1220
Lys Asn Ile Phe Ala Gln Leu Phe Ser Ala Arg Ile Ile Cys Ile Pro	
200 205 210	
ccg ttg caa gtg cta ttc ttt tcc aac tgt atc act cat gcg gtt ccg	1268
Pro Leu Gln Val Leu Phe Phe Ser Asn Cys Ile Thr His Ala Val Pro	
215 220 225	
gca gtc atg aac atc gga aga ctc cgc gac cca gcc aat tct ctt act	1316
Ala Val Met Asn Ile Gly Arg Leu Arg Asp Pro Ala Asn Ser Leu Thr	
230 235 240	
aaa aga gct gag aag tgg ctt ctt gaa cta gac gag cga acc cca cga	1364
Lys Arg Ala Glu Lys Trp Leu Leu Glu Leu Asp Glu Arg Thr Pro Arg	

245	250	255	260	
gcc gag aag ggc ttt ttc ttt tat ggt gaa gga tcc aac act tac gtt Ala Glu Lys Gly Phe Phe Phe Tyr Gly Glu Gly Ser Asn Thr Tyr Val 265 270 275				1412
tgc aac gtc caa gag caa ata gac cac gaa cgc cgg aag gtt gcc gca Cys Asn Val Gln Glu Gln Ile Asp His Glu Arg Arg Lys Val Ala Ala 280 285 290				1460
gcg tgt gga ttg cgt ctc aat tct ctc ttg cag gaa tgc aat gat gaa Ala Cys Gly Leu Arg Leu Asn Ser Leu Leu Gln Glu Cys Asn Asp Glu 295 300 305				1508
tat gat act gac tat gaa act ttg agg gaa tac tgc cta gca ccg tca Tyr Asp Thr Asp Tyr Glu Thr Leu Arg Glu Tyr Cys Leu Ala Pro Ser 310 315 320				1556
cct cat aac gtg cat cat gca tgc cct gac aac atg gaa cat cgc tat Pro His Asn Val His His Ala Cys Pro Asp Asn Met Glu His Arg Tyr 325 330 335 340				1604
ttt tct gaa gaa tta tgc tcg ttg gag gat gtc gcg gca att gca gct Phe Ser Glu Glu Leu Cys Ser Leu Glu Asp Val Ala Ala Ile Ala Ala 345 350 355				1652
att gcc aac atc gaa cta ccc ctc acg cat gca ttc atc aat att att Ile Ala Asn Ile Glu Leu Pro Leu Thr His Ala Phe Ile Asn Ile Ile 360 365 370				1700
cat gcg ggg aaa ggc aag att aat cca act ggc aaa tca tcc agc gtg His Ala Gly Lys Gly Lys Ile Asn Pro Thr Gly Lys Ser Ser Ser Val 375 380 385				1748
att ggt aac ttc agt tcc agc gac ttg att cgt ttt ggt gct acc cac Ile Gly Asn Phe Ser Ser Ser Asp Leu Ile Arg Phe Gly Ala Thr His 390 395 400				1796
gtt ttc aat aag gac gag atg gtg gag taa agaaggagtg cgtcgaagca Val Phe Asn Lys Asp Glu Met Val Glu 405 410				1846
gatcgttcaa acatttggca ataaagtttc ttaagattga atcctgttgc cgggtcttgcg				1906
atgattatca tataatttct gttgaattac gttaagcatg taataattaa catgtaatgc				1966
atgacgttat ttatgagatg gggtttttatg attagagtcc cgcaattata catttaatac				2026
gcgatagaaa acaaaatata gcgcgcaaac taggataaat tatcgcgcgcg ggtgtcatct				2086
atgttactag atcgatcaaa cttcgggtact gtgtaatgac gatgagcaat cgagaggctg				2146
actaacaaaa ggtatgccca aaaacaacct ctccaaactg tttcgaattg gaagtttctg				2206
ctcatgccga caggcataac ttagatatcc gcgggctatt ccactaatt cgtcctgctg				2266
gtttgcgcca agataaatca gtgcatctcc ttacaagttc ctctgtcttg tgaaatgaac				2326
tgctgactgc cccccaagaa agcctcctca tctcccagtt ggcgggcggt gatacaccat				2386
cgaaaaccca cgtccgaaca cttgatacat gtgcctgaga aataggccta cgtccaagag				2446
caagtccttt ctgtgctcgt cggaattcc tctcctgtca gacggtcgtg cgcatgtctt				2506

gcgttgatga agctt

2521

<210> 17  
<211> 413  
<212> PRT  
<213> Agrobacterium tumefaciens

<400> 17

Met Ala Ile Thr Leu Ser Ala Thr Ser Leu Pro Ile Ser Ala Ala Asp  
1 5 10 15

His His Pro Leu Pro Leu Thr Val Gly Val Leu Gly Ser Gly His Ala  
20 25 30

Gly Thr Ala Leu Ala Ala Trp Phe Ala Ser Arg His Val Pro Thr Ala  
35 40 45

Leu Trp Ala Pro Ala Asp His Pro Gly Ser Ile Ser Ala Ile Lys Ala  
50 55 60

Asn Glu Gly Val Ile Thr Thr Glu Gly Met Ile Asn Gly Pro Phe Arg  
65 70 75 80

Val Ser Ala Cys Asp Asp Leu Ala Ala Val Ile Arg Ser Ser Arg Val  
85 90 95

Leu Ile Ile Val Thr Arg Ala Asp Val His Asp Ser Phe Val Asn Glu  
100 105 110

Leu Ala Asn Phe Asn Gly Glu Leu Ala Thr Lys Asp Ile Val Val Val  
115 120 125

Cys Gly His Gly Phe Ser Ile Lys Tyr Glu Arg Gln Leu Arg Phe Lys  
130 135 140

Arg Ile Phe Glu Thr Asp Asn Ser Pro Ile Thr Ser Lys Leu Ser Asp  
145 150 155 160

Gln Lys Lys Cys Asn Val Asn Ile Lys Glu Met Lys Ala Ser Phe Gly  
165 170 175

Leu Ser Cys Phe Pro Ile His Arg Asp Asp Ala Gly Val Ile Asp Leu  
180 185 190

Pro Glu Asp Thr Lys Asn Ile Phe Ala Gln Leu Phe Ser Ala Arg Ile  
195 200 205

Ile Cys Ile Pro Pro Leu Gln Val Leu Phe Phe Ser Asn Cys Ile Thr  
210 215 220

His Ala Val Pro Ala Val Met Asn Ile Gly Arg Leu Arg Asp Pro Ala  
225 230 235 240

Asn Ser Leu Thr Lys Arg Ala Glu Lys Trp Leu Leu Glu Leu Asp Glu  
245 250 255

Arg Thr Pro Arg Ala Glu Lys Gly Phe Phe Phe Tyr Gly Glu Gly Ser  
260 265 270

Asn Thr Tyr Val Cys Asn Val Gln Glu Gln Ile Asp His Glu Arg Arg  
275 280 285

Lys Val Ala Ala Ala Cys Gly Leu Arg Leu Asn Ser Leu Leu Gln Glu  
290 295 300

Cys Asn Asp Glu Tyr Asp Thr Asp Tyr Glu Thr Leu Arg Glu Tyr Cys  
305 310 315 320

Leu Ala Pro Ser Pro His Asn Val His His Ala Cys Pro Asp Asn Met  
325 330 335

Glu His Arg Tyr Phe Ser Glu Glu Leu Cys Ser Leu Glu Asp Val Ala  
340 345 350

Ala Ile Ala Ala Ile Ala Asn Ile Glu Leu Pro Leu Thr His Ala Phe  
355 360 365

Ile Asn Ile Ile His Ala Gly Lys Gly Lys Ile Asn Pro Thr Gly Lys  
370 375 380

Ser Ser Ser Val Ile Gly Asn Phe Ser Ser Ser Asp Leu Ile Arg Phe  
385 390 395 400

Gly Ala Thr His Val Phe Asn Lys Asp Glu Met Val Glu  
405 410

<210> 18

<211> 835

<212> DNA

<213> Streptomyces hygroscopicus

<400> 18

gctcgctgtc attttcgaga cgccatcttt ggaagcgggtg gccgaatccg tactgcgcgg 60

actcgacgac gcgtaaaacg atcgaccacg tacacgagtc cggacacggg gcgaggaggc 120

ccggttccgg caccgaggaa gaccgaagga agaccacacg tgagcccaga acgacgcccg 180

gccgacatcc gccgtgccac cgaggcggac atgccggcgg tctgcaccat cgtcaaccac 240



tacatcgaga	caagcacggt	caacttccgt	accgagccgc	aggaaccgca	ggagtggacg	300
gacgacctcg	tccgtctgcg	ggagcgctat	ccctggctcg	tcgccgaggt	ggacggcgag	360
gtcgccggca	tcgcctacgc	gggcccctgg	aaggcacgca	acgcctacga	ctggacggcc	420
gagtcgaccg	tgtacgtctc	cccccgccac	cagcggacgg	gactgggctc	cacgtcttac	480
accacactgc	tgaagtccct	ggaggcacag	ggcttcaaga	gcgtggctcg	tgtcatcggg	540
ctgcccacg	acccgagcgt	gcgcatgcac	gaggcgctcg	gatatgcccc	ccgcggcgatg	600
ctgcggggcg	ccggcttcaa	gcacgggaac	tggcatgacg	tgggtttctg	gcagctggac	660
ttcagcctgc	cggtagccgc	ccgtccggtc	ctgcccgta	ccgagatctg	aacggagtgc	720
gcgtgggcat	cgcgcgagtt	ggagctggta	cgggaactca	tcgaactcaa	ctggcataacc	780
cgcaatggtg	aggtggaacc	gcggcggtac	gcgtacgacc	gtgccaggga	ggcct	835

<210> 19  
 <211> 623  
 <212> DNA  
 <213> *Oryza sativa*

<400> 19						
tatatacata	cccccccctc	tcttcccatc	cccccaaccc	taccaccacc	accaccacca	60
cctctctccc	cctcgctgcc	ggacgacgag	ctctctcccc	ctccccctcc	gccgccgccg	120
gtaaccaccc	cgcgtccctc	tctcttttct	ttctccgttt	tttttttccg	tctcgctctcg	180
atctttggcc	ttggtagttt	gggggcgaga	ggcggcttcg	tcgccagat	cggtagcgcg	240
gagggggcgg	atctcgcgcc	tgggtctcgg	cgtgcggccg	gatcctcgcg	gggaatgggg	300
ctctcgcatg	tagatctgat	ccgccgttgt	tgggggagat	gatggggcgt	ttaaaatttc	360
gccatgctaa	acaagatcag	gaagagggga	aaagggcact	atggtttata	tttttatata	420
tttctgctgc	tgctcgtcag	gcttagatgt	gctagatctt	tctttcttct	ttttgtgggt	480
agaatttgaa	tccctcagca	ttgttcatcg	gtagtttttc	ttttcatgat	ttgtgacaaa	540
tgcagcctcg	tgcggagctt	ttttgtaggt	agaagatggc	tgacgccgag	gatgggggat	600
ccccgggtgg	tcagtccctt	atg				623

<210> 20  
 <211> 1506  
 <212> DNA  
 <213> *Trichoderma longibrachiatum*

<400> 20						
accatggctc	aatctgcttg	tactcttcaa	tctgagactc	atcctccact	tacttggcag	60
aagtgttcat	ctggtggtac	ttgtactcaa	cagactggat	ctgttggtat	tgatgctaac	120
tggagatgga	ctcatgctac	taactcttct	actaactgct	atgatggtaa	cacttgggtca	180

tctactcttt	gtcctgataa	cgagacttgt	gctaagaact	gctgtcttga	tggtgctgct	240
tacgcttcta	cttacggagt	tactacttct	ggaaactctc	tttctattgg	attcgttact	300
cagtctgctc	agaagaacgt	tggtgctagg	ttgtacttga	tggtctctga	tactacttac	360
caagagttca	ctcttcttgg	taacgagttc	tctttcgatg	ttgatgtttc	tcaacttcca	420
tgtggtttga	acggagcttt	gtacttcggt	tctatggatg	ctgatgggtg	agtttctaag	480
tatccaacta	acactgctgg	agctaagtat	ggtactgggt	actgtgattc	tcagtgtcca	540
agagatctta	agttcattaa	cggacaagct	aatgttgagg	gatgggagcc	atcttctaac	600
aatgctaaca	ctggtattgg	aggtcatgga	tcttgttgct	ctgagatgga	tatttgggag	660
gctaactcta	tctctgaggc	tcttactcca	catccatgca	ctactgttgg	acaagagatt	720
tgcgagggag	atggttggtg	tggaacttac	tctgataaca	gatacggagg	tacttgcgat	780
ccagatggat	gtgattggaa	tccatacaga	cttggttaaca	cttctttcta	cgggtccagga	840
tcttcattca	ctcttgatac	tactaagaag	ttgactgttg	ttactcagtt	cgagacttct	900
ggtgctatca	acagatacta	cgttcagaat	ggagttaact	tccaacaacc	taacgctgag	960
cttggttctt	actctggtaa	cgagttgaac	gatgattact	gtactgctga	ggaagctgag	1020
tttggtggat	catctttctc	tgataagggg	ggacttactc	agttcaagaa	ggctacttct	1080
ggtggtatgg	ttcttgttat	gtctctttgg	gatgattact	acgctaacat	gttgtggctt	1140
gattctactt	acccaactaa	cgagacttct	tctactccag	gtgctgttag	aggatcttgc	1200
tctacttctt	ctggtgttcc	tgctcaagtt	gaatctcaat	ctcctaatgc	taaggttact	1260
ttctctaaca	tcaagttcgg	accaattgga	tctactggta	acccttctgg	aggtaatcca	1320
cctggaggta	atccacctgg	aactactaca	actaggagac	cagctactac	aactggatca	1380
tctccaggac	ctactcaatc	tcattacggg	caatgtggag	gtattggtta	ctctgggtcca	1440
actgtttgtg	cttctggaac	tacttgtcaa	gttcttaacc	cttactattc	tcaatgcctt	1500
taatga						1506

<210>	21
<211>	26
<212>	DNA
<213>	synthetic DNA

<400> 21

<210> 22

<211> 25  
<212> DNA  
<213> synthetic DNA

<220>  
<221> misc\_feature  
<222> (1)..(25)  
<223> PCR primer SP3R

<400> 22

cgctgtacac gcacctgac ctgcc

25

CGCTGTACAC GCACCTGAC CTGCC